Nonparametric approach to analysing operational risk losses

Universitat de Barcelona
Risk in Finance and Insurance Research Group

Catalina Bolancé, Mercedes Ayuso and Montserrat Guillén

www.ub.edu/rfa
Motivation

Operational Risk

Fraud detection and fraud control

Methodology

The Data

Results

Conclusions
Motivation

- To analyze **fraud as an operational risk** for the insurance company.
- To quantify the **effect of failing to detect fraud** in a real case study.
- To apply a **nonparametric approach** to measure operational risk due to claims fraud.
Operational Risk

According to Article 101 of the Solvency II Level 1 text:
- The necessary capital shall correspond to the Value-at-Risk of the basic own funds of an insurance or reinsurance undertaking subject to a confidence level of 99.5% over a one-year period.

According to Article 13(33) of the Solvency II Level 1 text:
- **Operational Risk** is the risk of loss arising from inadequate or failed internal processes, or from personnel and systems, or from external events. It refers to the possibility of unexpected events that occur as a consequence of alterations in regular functioning of companies.
Operational Risk

Difficulties:

- In insurance some events have **no physical consequences**, and therefore they are more difficult to recognize than operational failures elsewhere.
- **The lack of data.**
  - Underreporting (*undetected fraud*)
  - Large operational events are infrequent (*do insurers mainly focus on large fraudulent claims?*)
Operational Risk

- Undetected fraudulent claims is a type of operational failure.
- The absence or inefficiency of an auditing protocol can lead the insurance firm to severe aggregate losses.
- Investing in an auditing system reduces the risk of losses in the operational risk category, ....however the classical approach to fraud detection is usually based on a cost-benefit analysis.

Literature:
Fraud detection and fraud control

Different areas associated with fraud as an operational risk

**Figure:** Sources of fraud for the insurance company
Fraud detection and fraud control

When aiming to detect external fraud arising from incoming claims, two possible courses of action are possible:

1. Assume that fraud exists but the insurer has no active policy to detect fraudulent claims or to prevent fraud.

2. Set up specialized departments to fight against fraud (SIUs, Special Investigation Units).
Fraud detection and fraud control

Figure: Claims processing and auditing (Viaene et al., 2007)
**Methodology**

We simulate total operational risk costs in fraud detection. The procedure is:

1. We fit the cumulative distribution function (cdf) $F_X$ of the random variable $X$, which is equal to cost of a single claim associated to an operational loss. We use **the double transformed kernel estimation** of cdf.

2. We fix the **proportion of fraudulent claims** in the motor insurance claims database. We suppose:
   - Number of claims $N_{Fr}$ can be deterministic.
   - Number of claims $N_{Fr}$ can be stochastic.

3. We simulate the cost of the $N_{Fr}$ claims ($x = F_X^{-1}(u)$) using the previously fitted cdf and the sum all simulated costs. We do 10,000 replications. Then we obtain the **distribution of the total loss** due to fraudulent claims.
Methodology

Double transformed kernel estimation of cdf

\[
\hat{F}_X(x) = \frac{1}{n} \sum_{i=1}^{n} K \left( \frac{M^{-1}(T(x)) - M^{-1}(T(X_i))}{b} \right)
\]

where \( T \) is the cdf of the Generalized Champernowne distribution:

\[
T(x) = \frac{(x + c)^\alpha - c\alpha}{(x + c)^\alpha + (M + c)^\alpha - 2c\alpha}
\]  \hspace{1cm} (1)

and \( M \) is the cdf of \( \text{Beta}(3, 3) \) in the domain \((-1, 1)\):

\[
M(x) = \frac{3}{16} x^5 - \frac{5}{8} x^3 + \frac{15}{16} x + \frac{1}{2}, \ x \in (-1, 1).
\]  \hspace{1cm} (2)

Bolancé and Guillén (2011 WP), Bolancé, Guillén and Nielsen (2003, IME, 2008 Stat&Prob Letters) and Buch-Larsen et al. (2005 Statistics)
Methodology

Double transformed kernel estimation of cdf

The asymptotic mean square error (MSE) of the transformed kernel estimator or double transformed kernel estimator of cdf is:

\[
\frac{F_X(x)[1-F_X(x)]}{n} - \frac{1}{T'(x)} f_X(x) \frac{b}{n} \left( 1 - \int_{-1}^{1} K^2(t) \, dt \right) \\
+ \frac{1}{T''(x)} \left( 1 - \frac{f_X(x)}{T'(x)} \right)^2 \left[ \frac{1}{2} f_X(x) \int_{-1}^{1} t^2 k(t) \, dt \right]^2 b^4.
\]

The first two terms in (3) correspond to the variance and the last term is the squared bias. The variance part in (3) is smaller than the variance of classical kernel estimator if \( T'(x) < 1 \) and, on the contrary, the squared bias is larger than the bias of the classical kernel estimator.
The Data

- We use a sample taken from a *motor insurance portfolio* which was collected by a Spanish insurer and which contains information of 64,642 claims.

- The claims are about *property damages to the vehicle* and were reported during 2000; 691 claims are fraudulent, and the rest are not fraudulent.

- We assume those that are not found to be fraudulent could still contain some form of *undetected fraud*. 
### The Data

**Table: Descriptive statistics of claims costs (in Euros)**

<table>
<thead>
<tr>
<th></th>
<th>Non fraudulent claims (N=63,951)</th>
<th>Fraudulent claims (N=691)</th>
<th>All claims (N=64,642)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average</td>
<td>Std. Dev.</td>
<td>Minimum</td>
</tr>
<tr>
<td>Claimed property damages (A)</td>
<td>612.45</td>
<td>1,178.79</td>
<td>0.00</td>
</tr>
<tr>
<td>Auditing cost (B)</td>
<td>38.94</td>
<td>26.65</td>
<td>0.00</td>
</tr>
<tr>
<td>Final compensation (C)</td>
<td>612.45</td>
<td>1,178.79</td>
<td>0.00</td>
</tr>
<tr>
<td>Total cost of claim (B+C)</td>
<td>651.39</td>
<td>1,199.02</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*Only in 8 cases the final compensation to the insured was different from zero*
Results

- We estimate the 99.5% quantile of the total cost of undetected fraud.
- The percentage of undetected fraudulent claims in the insurance portfolio is thought to be around 5%, even if some special audit exists.
- We have set a series of scenario levels ranging from 0.1% to 10% for the proportion of fraudulent claims.
- We compare normal fit with nonparametric fit (double transformed estimation) and we consider deterministic and stochastic number of claims $N_{Fr}$.
- We calculate the auditing costs in the different scenarios of proportion of fraudulent claims.
Table: Estimated 99.5% Quantile for Operational Risk (in thousand Euros)

<table>
<thead>
<tr>
<th>Fraud Frequency</th>
<th>$N_{Fr}$ deterministic</th>
<th>Normal fit VaR$_{0.995}$</th>
<th>Nonparametric fit VaR$_{0.995}$</th>
<th>Nonparametric fit $N_{Fr}$ stochastic VaR$_{0.995}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1%</td>
<td>102.99</td>
<td>560.47</td>
<td>663.62</td>
<td></td>
</tr>
<tr>
<td>0.3%</td>
<td>269.93</td>
<td>1,581.71</td>
<td>1,737.66</td>
<td></td>
</tr>
<tr>
<td>0.5%</td>
<td>429.85</td>
<td>2,590.33</td>
<td>2,802.31</td>
<td></td>
</tr>
<tr>
<td>1%</td>
<td>819.37</td>
<td>5,059.18</td>
<td>5,371.40</td>
<td></td>
</tr>
<tr>
<td>3%</td>
<td>2,334.68</td>
<td>14,824.58</td>
<td>15,368.61</td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>3,827.80</td>
<td>24,555.51</td>
<td>25,168.01</td>
<td></td>
</tr>
<tr>
<td>6%</td>
<td>4,570.60</td>
<td>29,387.41</td>
<td>30,098.21</td>
<td></td>
</tr>
<tr>
<td>7%</td>
<td>5,311.73</td>
<td>34,208.80</td>
<td>34,947.18</td>
<td></td>
</tr>
<tr>
<td>8%</td>
<td>6,051.54</td>
<td>39,072.85</td>
<td>39,868.76</td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>7,528.09</td>
<td>48,701.35</td>
<td>49,780.99</td>
<td></td>
</tr>
</tbody>
</table>

$N_{Fr}$ is the number of fraudulent claims
## Results

### Table: Auditing cost (in thousand Euros)

<table>
<thead>
<tr>
<th>Fraud Frequency</th>
<th>Total auditing cost</th>
<th>Fraud auditing cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1%</td>
<td>2,636.23</td>
<td>14.28</td>
</tr>
<tr>
<td>0.3%</td>
<td>2,755.19</td>
<td>43.06</td>
</tr>
<tr>
<td>0.5%</td>
<td>2,874.33</td>
<td>72.06</td>
</tr>
<tr>
<td>1%</td>
<td>2,993.29</td>
<td>144.11</td>
</tr>
<tr>
<td>3%</td>
<td>3,112.43</td>
<td>432.57</td>
</tr>
<tr>
<td>5%</td>
<td>3,231.39</td>
<td>721.02</td>
</tr>
<tr>
<td>6%</td>
<td>3,350.35</td>
<td>865.13</td>
</tr>
<tr>
<td>7%</td>
<td>3,469.49</td>
<td>1,009.25</td>
</tr>
<tr>
<td>8%</td>
<td>3,588.45</td>
<td>1,153.58</td>
</tr>
<tr>
<td>10%</td>
<td>3,707.59</td>
<td>1,442.03</td>
</tr>
</tbody>
</table>
Conclusions

- The non-detection of fraudulent claims may be a source of operational risk within an insurance company.
- According to our results, **claims auditing and detection considerably reduces operational risk.**
- The relative gain depends on the proportion of fraud in the portfolio and the cost of detecting fraud.
Nonparametric approach to analysing operational risk losses

Universitat de Barcelona
Risk in Finance and Insurance Research Group

Catalina Bolancé, Mercedes Ayuso and Montserrat Guillén

www.ub.edu/rfa